JOINT MODELLING

OF

LONGITUDINAL MEASUREMENTS

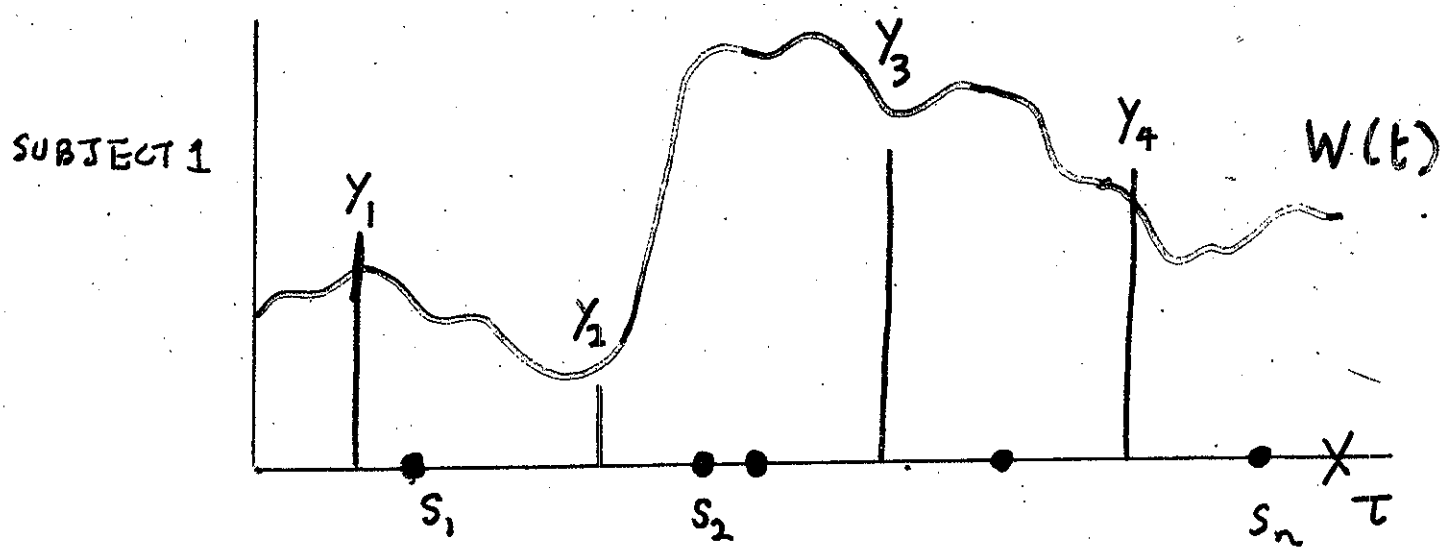
AND

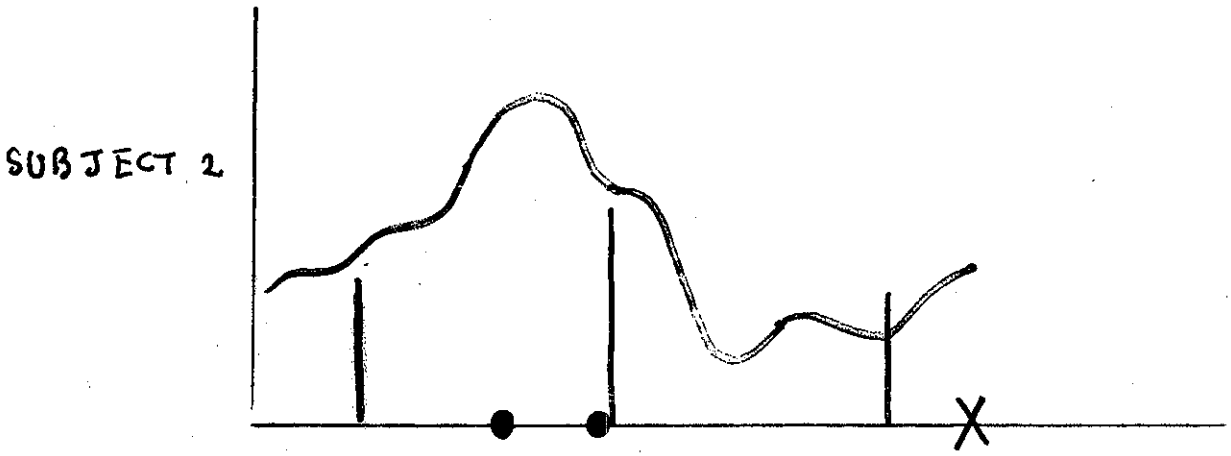
EVENT HISTORIES

- Peter Diggle (Lancaster)
- Rob Henderson (Lancaster)
- Jane Xu (Johns Hopkins)
- Scott Zeger (Johns Hopkins)

With thanks to ESRC, NATO, MRC, Janssen Pharmaceuticals.

Generic data format





Scientific interest may include:

- joint evolution of the measurement and event-time processes;
- adjustment of inferences about longitudinal measurements to allow for possibly outcome-dependent dropout;
- use of intermediate longitudinal measurements as surrogate for time to terminating event.

Example: schizophrenia trial

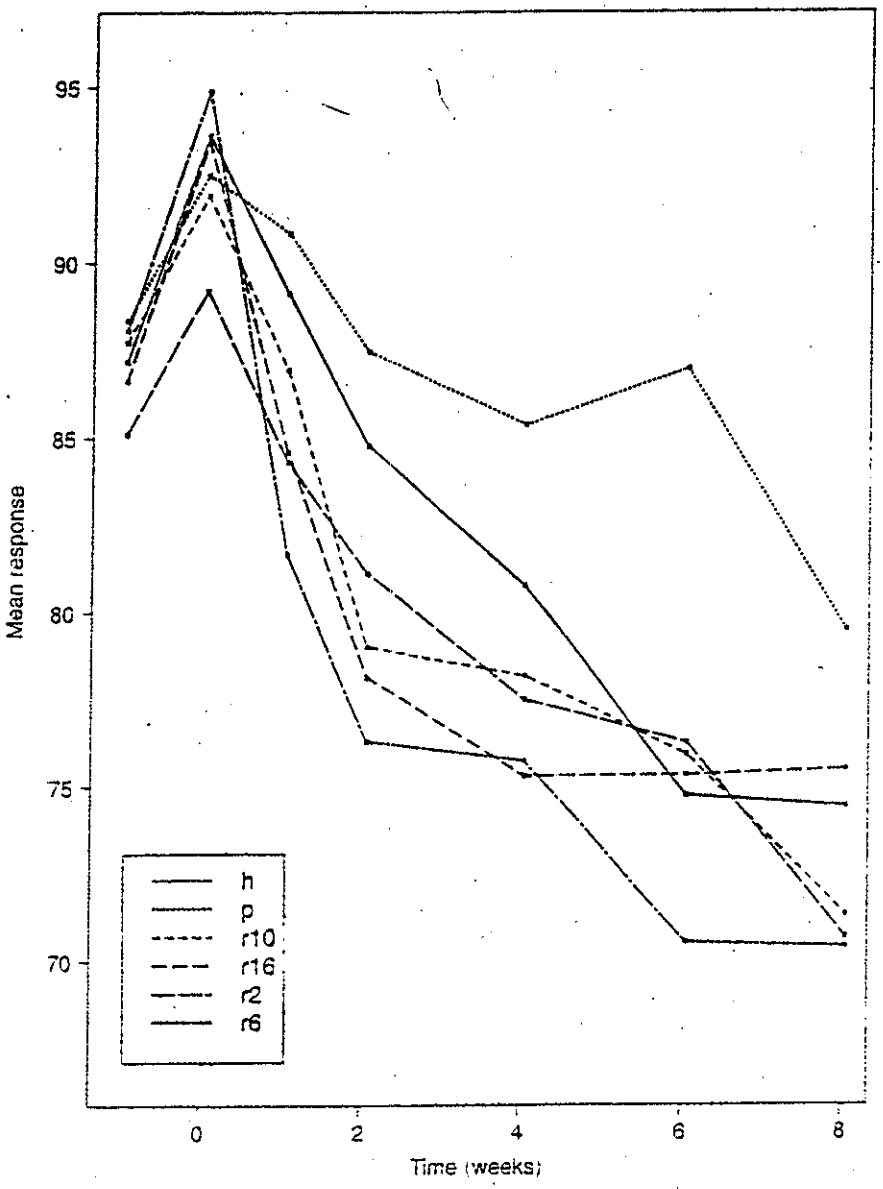
- multi-centre, double blind, parallel group study
- 523 patients, randomly allocated amongst six treatments:
  - placebo
  - haloperidol 20mg (standard therapy)
  - risperidone 2mg, 6mg, 10mg and 16mg (novel therapy)
- response variable was a measure of psychiatric disorder (PANSS)
- measurements intended to be taken at weeks:
  -1 (selection), 0 (baseline), 1, 2, 4, 6, 8
- 270 dropouts, for following stated reasons:

Abnormal lab result	4
Adverse experience	26
Inadequate response	183
Inter-current illness	3
Lost to follow-up	3
Other reason	7
Uncooperative	25
Withdrew consent	19

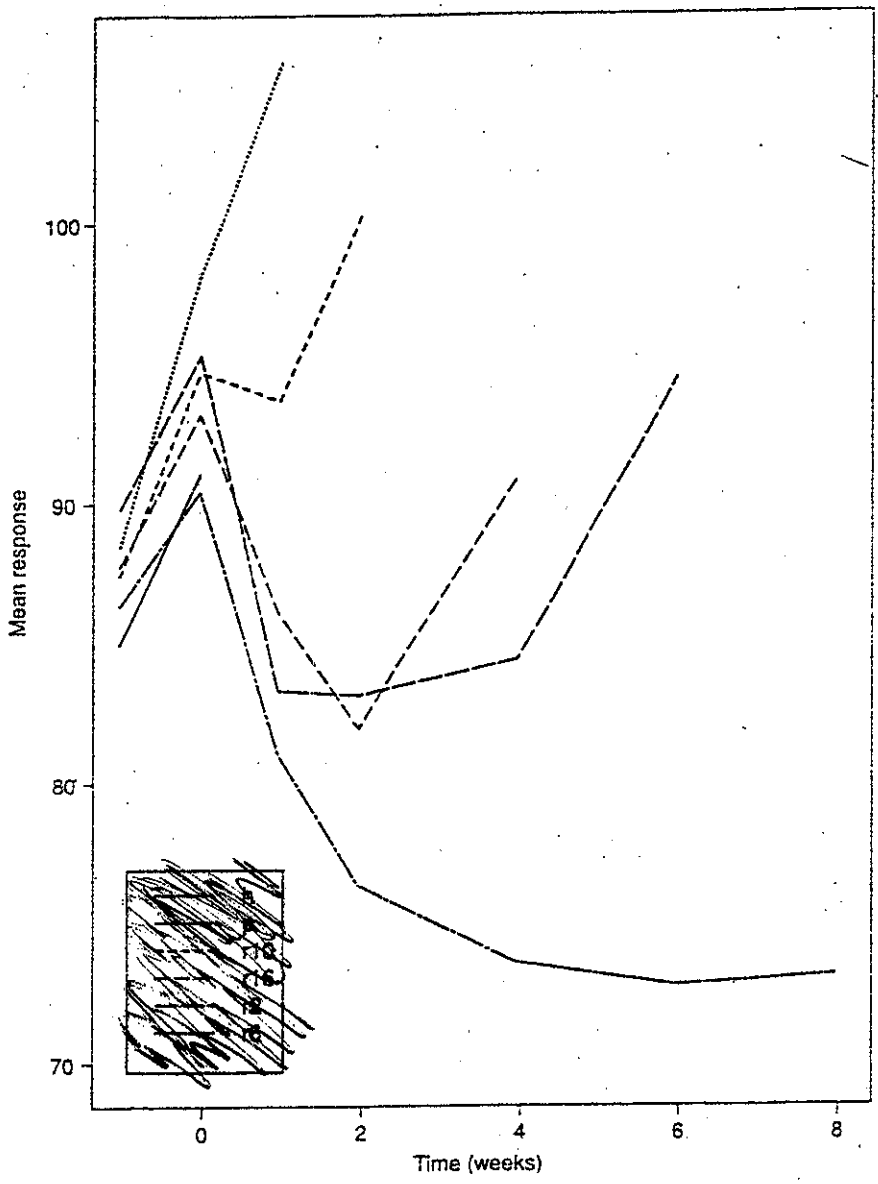
Clinical objective

achieve reduction of at least 20% in mean PANSS score

Observed mean response by time and treatment



Observed mean response by dropout cohort



Proposed class of models

1. Latent process

$W(t) = \{W_1(t), W_2(t)\}$ a bivariate Gaussian process

- $W_k(t) = D_k(t)U_k + V_k(t)$
- $\{V_1(t), V_2(t)\}$: bivariate stationary Gaussian process
- (U_1, U_2) : multivariate Gaussian random effects

$W(t)$ realised independently for each subject

2. Measurement model

$$Y_{ij} = \mu_i(t_{ij}) + W_{1i}(t_{ij}) + Z_{ij}$$

- $Z_{ij} \sim \mathrm{N}(0, \tau^2)$
- $\mu_i(t_{ij}) = X_{1i}(t_{ij})\beta_1$

3. Intensity model

$$\lambda_i(t) = R_i(t)\alpha_0(t)\mathcal{F}(X_{2i}(t)\beta_2 + W_{2i})$$

- $\alpha_0(t)$ = non-parametric baseline intensity
- $R_i(t)$ = "at risk" indicator
- typical choice for $\mathcal{F}$ might be $\mathcal{F}(\eta_2) = \exp\{W_2(t)\}$

4. Special case for preliminary analysis

$$W_2(t) = \gamma W_1(t)$$

Hence, γ is a single parameter which measures association between measurement process and event intensity.

Likelihood evaluation

Notation

Y : measurement data

W : (bivariate) latent process

N : event history data

- Conditional independence: $N \perp Y \mid W$
- Standard marginal for Y: $L_1(\theta, Y)$
- Easy conditional distribution $[W \mid Y]$
- Standard conditional for $[N \mid W]$: $L_2(\theta, N \mid W)$
- $\Rightarrow$ selection factorisation

$$L(\theta) = L_1(\theta) \times E_{W|Y}[L_2(\theta, N \mid W)]$$

Requires infinite-dimensional integration wrt W?

No – non-parametric specification for baseline hazard implies we only need W at event times

A score test for association

- joint analysis of Y and D computationally intensive but separate analyses straightforward
- hence, may be useful to conduct a preliminary test of association between Y and D
- score test is based on slope of log-likelihood at $H_0 : \gamma = 0$ and is therefore obtainable from separate analyses of Y and D

Resulting test statistic is

$$U = \sum_{i=1}^{m} \int_0^\tau E_{\mathcal{W}_1^S|Y}[W_{1i}(t)]dM_i(t)$$

where

$$M_i(t) = N_i(t) - \Lambda_i(t) = N_i(t) - \int_0^t R_i(u)e^{x_{2i}(u)'\beta_2}dA_0(u)$$

$N_i(t)$ = number of $s_{ij} \leq t$

$R_i(t)$ = "at risk" indicator

$\Lambda_i(t) = \int_0^t \lambda_i(s)ds$

$A_0(t) = \int_0^t \alpha_0(s)ds$

Properties of score test

Derive Normal approximation to null distribution of score test statistic using either:

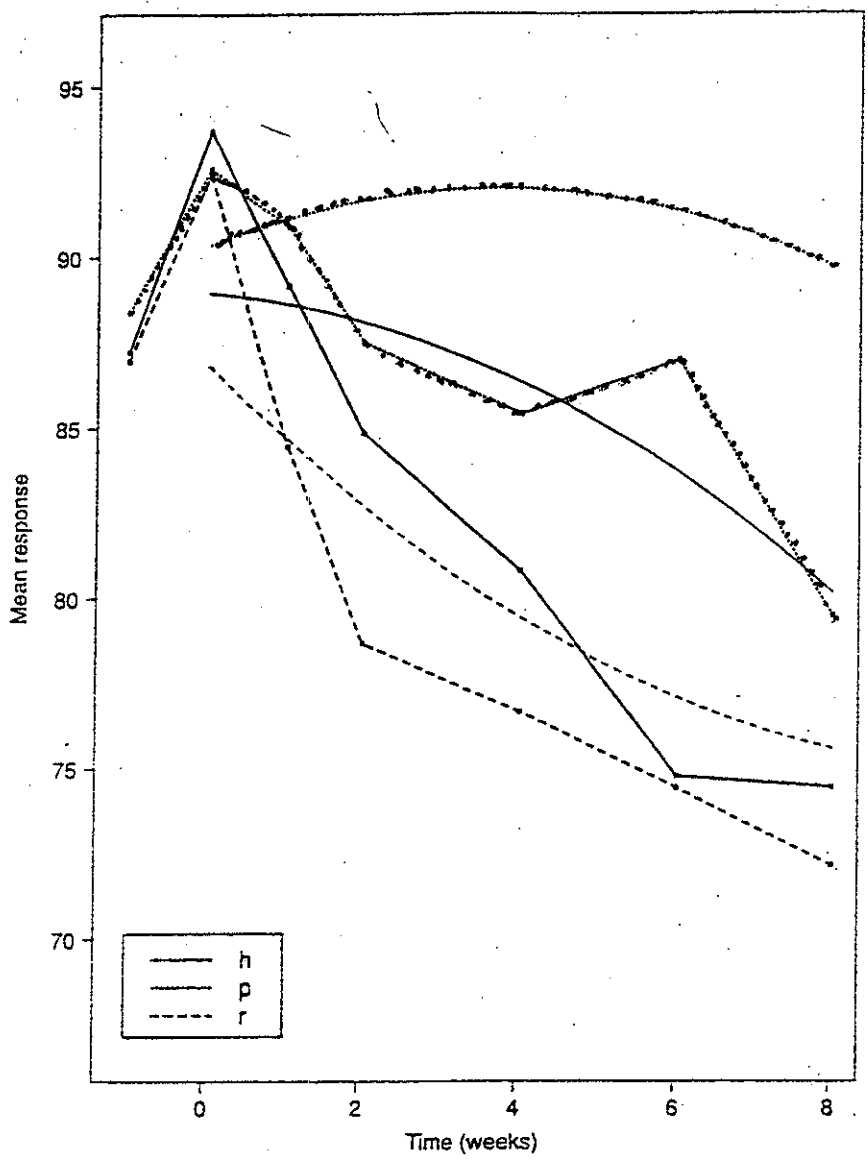
- direct calculation (test statistic as linear functional)
- Martingale central limit theorem (simpler, and gives similar result)

Simulation study with sample sizes comparable to risperidone trial suggests:

- nominal size OK
- power increases with γ
- power increases with strength of serial correlation

Application to risperidone trial

Observed and fitted means



Score test: $N(0,1) = 9.86$

What if we include adjustment for baseline?

Score test: $N(0,1) = 9.30$

Estimation in presence of association

$$L(\theta) = L_1(\theta) \times E_{W|Y}[L_2(\theta, N \mid W)]$$

3. Intensity model re-visited

$$\lambda_i(t) = R_i(t)\alpha_0(t)\mathcal{F}(X_{2i}(t)\beta_2 + W_{2i})$$

- $\alpha_0(t)$ = non-parametric baseline intensity
- $R_i(t)$ = "at risk" indicator
- $\mathcal{F}(\eta_2) = \exp\{W_2(t)\}$

Options include:

- adopt fully parametric approach and use MCMC
- two-stage plug-in method
- non-iterative Monte Carlo evaluation
- quasi-EM

Two-stage plug-in method

- Replace $E_{W|Y}[L_2(\theta, N \mid W)$ with $L_2(\theta, N \mid E_{W|Y}[W])$
- Use partial likelihood PL_2 in place of L_2
- Maximise

$$L(\theta) = L_1(\theta) \times PL_2(\theta, N \mid E_{W|Y}[W])$$

Simulation experiment:

	Mean $\hat{\gamma}$				
Method	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.25$	$\gamma = 0.5$	
$E[W]$	**0.01**	**0.08**	**0.23**	**0.43**	$sd \simeq 0.08$

Non-iterative Monte Carlo evaluation

- Likelihood/partial likelihood

$$L(\theta) = L_1(\theta) \times E_{W|Y}[PL_2(\theta, N \mid W)]$$

- Estimate $E[PL_2]$ by Monte Carlo Integration
- More stable to estimate $E[\log PL_2]$?

Simulation experiment:

	Mean $\hat{\gamma}$				
Method	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.25$	$\gamma = 0.5$	
$E[W]$	0.01	0.08	0.23	0.43	$sd \simeq 0.08$
$\hat{E}[PL_2]$	**-0.01**	**0.07**	**0.18**	**0.38**	$sd \simeq 0.08$
$\hat{E}[\log PL_2]$	**-0.02**	**0.07**	**0.15**	**0.32**	$sd \simeq 0.06$

Quasi-EM

- W occurs in PL_2 only through $\exp(W)$
- EM algorithm: replace with $E_{W|(Y,N)}[\exp(W)]$
- Quasi-EM : use $E_{W|Y}[\exp(W)]$

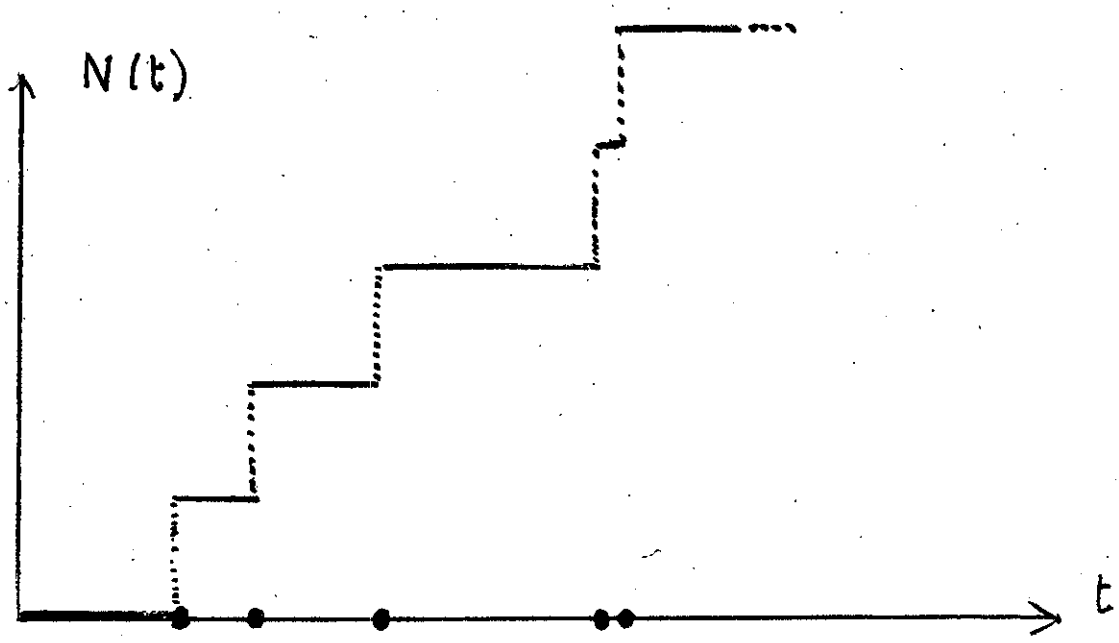
Simulation experiment:

	Mean $\hat{\gamma}$				
	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.25$	$\gamma = 0.5$	
$E[W]$	0.01	0.08	0.23	0.43	$sd \simeq 0.08$
$\hat{E}[PL_2]$	-0.01	0.07	0.18	0.38	$sd \simeq 0.08$
$\hat{E}[\log PL_2]$	-0.02	0.07	0.15	0.32	$sd \simeq 0.06$
$E[\exp(W)]$	**0.01**	**0.10**	**0.26**	**0.51**	$sd \simeq 0.09$

Current work

- develop analogy with omitted frailty in survival modelling
- quantify difference between $[W|Y, N]$ and $[W|Y]$
- extend modelling framework – under quasi-EM estimation, no particular advantage in restriction to $W_2 = \gamma W_1$

Martingale theory (RH)



$N(t)$ is a **counting process**, with conditional intensity $\lambda(t)$ such that

$$\mathrm{E}[dN(t)|\mathcal{F}_t] = \lambda(t)dt$$

where $\mathcal{F}_t$ denotes history of $N(t)$ up to $t-$

Cumulative intensity is

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

Then,

$$M(t) = N(t) - \Lambda(t)$$

is a **martingale**, with essential property that

$$\mathrm{E}[M(t)|\mathcal{F}_t] = M(t-)$$

Some properties

- **Martingale central limit theorem**: for large m,

$$m^{-\frac{1}{2}} \sum_{i=1}^{m} M_i(t) \sim \mathrm{N}(0, v(t))$$

  for known $v(t)$

- If $M(t)$ is a martingale and $h(t)$ is any left-continuous function then

$$H(t) = \int_0^t h(s) dM(s)$$

  is also a martingale, with variance

$$\int_0^t \{h(s)\}^2 d\Lambda(s)$$

Alternative variance calculation

$$V_2(\tau) = \sum_{i=1}^{m} \Big\{ \int_0^\tau E_{\mathcal{W}_1^S|Y}[W_{1i}^2(t)] d\Lambda_i(t) - \int_0^\tau \int_0^\tau \mathrm{Cov}_{\mathcal{W}_1^S|Y}(W_{1i}(t), W_{1i}(s)) dM_i(t) dM_i(s) \Big\}$$

Power study

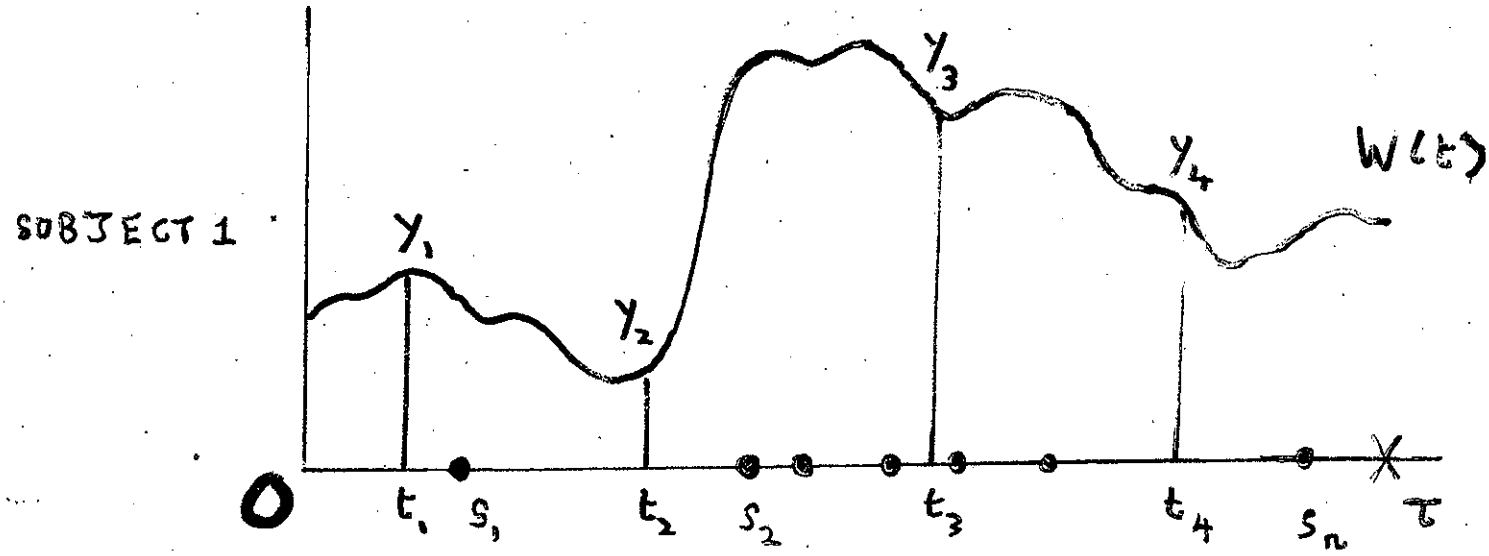
Simulation model:

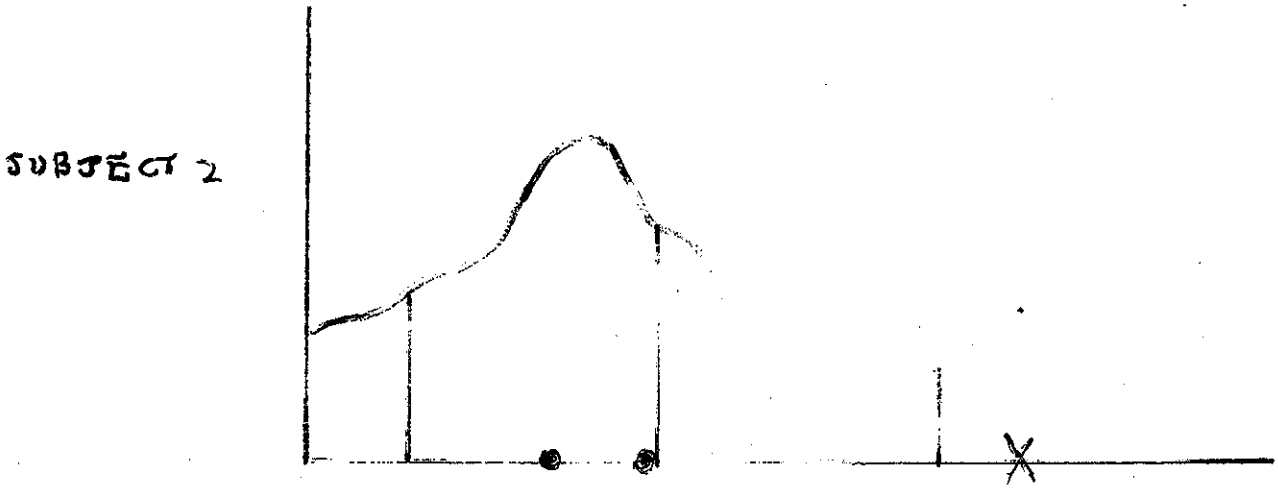
- $m = 250$ subjects in single group
- up to 4 measurements at times $t = 0, 10, 20, 30$, censoring time $\tau = 50$
- $\mu(t) = 5 + 0.1t$, $\sigma_z^2 = 0.25$
- $W_1(t) = U + V(t)$ where:
  - $U \sim \mathrm{N}(0, \sigma_u^2$
  - $V(t) \sim \mathrm{SGP}(0, \sigma_1^2, \exp(-|k|/\phi))$
  - ϕ such that lag-10 correlation is 0.5 or 0.05
  - $\sigma_u^2 + \sigma_1^2 = 1$

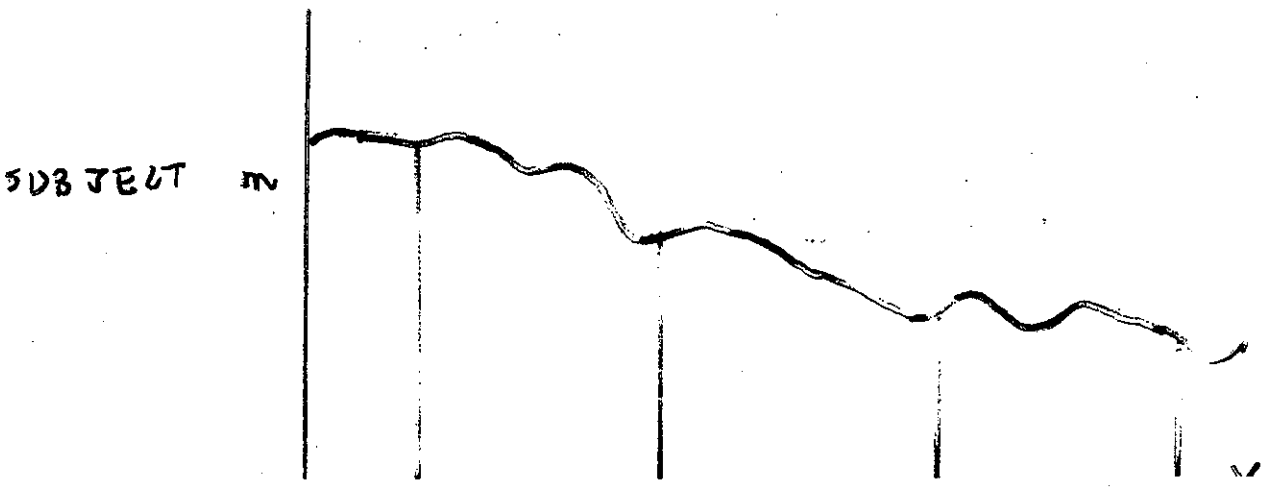
σ_u^2	σ_1^2	$\rho_1(10)$	$\sigma_u^2 + \sigma_1^2\rho_1(10)$	$\gamma = 0$	$\gamma = 0.1$	$\gamma = 0.25$	$\gamma = 0.5$
0.5	0.5	0.5	0.75	0.04	0.22	0.90	1.00
		0.05	0.525	0.04	0.10	0.67	1.00
0.8	0.2	0.5	0.90	0.08	0.21	0.92	1.00
		0.05	0.81	0.06	0.20	0.86	1.00
0.2	0.8	0.5	0.60	0.06	0.18	0.74	1.00
		0.05	0.24	0.06	0.08	0.36	0.94

- nominal size OK
- power increases with γ
- power increases with strength of serial correlation

Data format


SUBJECT 1
Y_1
Y_2
Y_3
Y_4
W(t)
0
t_1 s_1
t_2 s_2
t_3
t_4
s_n
τ


SUBJECT 2


SUBJECT m